PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
G01N 33/53, C12Q 1/68

A1

(11) International Publication Number: WO 99/06833

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/US98/15895

(22) International Filing Date: 29 July 1998 (29.07.98)

(30) Priority Data:
08/903,035 29 July 1997 (29.07.97) US
08/903,035 14 May 1998 (14.05.98) US

(71)(72) Applicants and Inventors: HUSTON, James, S. [US/US]; 5 Drew Road, Chestnut Hill, MA 02167 (US). HUTCHENS, T., William [US/US]; 2109 Alamede Avenue, David, CA 95616 (US).

(74) Agents: LOREN, Ralph, A. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHODS FOR OBTAINING PROBES OF PEPTIDE FUNCTION

(57) Abstract

Methods for obtaining probes of function for peptides are described. The methods include contacting an array of peptidic targets with a diverse library of binding partners, and identifying the binding partners which bind to the array of targets.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHODS FOR OBTAINING PROBES OF PEPTIDE FUNCTION

Background of the Invention

The ability of modern DNA sequencing techniques to produce large amounts of DNA sequence information has increased in recent years. Undertakings such as the human genome project, and other projects to sequence the entire genomes of bacteria, plants and lower animals, are producing DNA sequence information at a rapid and accelerating rate. Completion of sequencing of the human genome, if accomplished over the next few years, requires that in each year of this effort new sequence data and analysis must be produced at the rate of 500 million base pairs per year, i.e., the equivalent of the *E. coli* genome being published every 2 days for approximately 7 years (Adams, M.D. and J.C. Venter, 1996 Science 274:534-6). Preliminary to sequencing this genome, efforts at mapping restriction fragment length polymorphisms (RFLPs), and acquiring a library of expressed sequence tags (ESTs) representative of the entire set of human genes, have been largely successful, resulting in publication of a gene map of the human genome (Schuler, G.D. et al, 1996 Science 274:540-546). However, the proteins which are encoded by the sequenced DNA are not always known, and cloning of large numbers of genes may not always be practical.

Alternative approaches to determining the function of genes, including: hybridization studies with RNA to determine site of gene activity; cataloging human variants of a gene; studies of gene function in simpler organisms (yeast, nematodes, and fruit flies), including production and analysis of deletion mutations for all identified genes; identification of the entire complement of cellular proteins; and systematic catalogs of protein interactions (Lander, E. 1996 Science 274:536-539), have also been proposed. However, additional methods for determining the functions of proteins encoded by gene sequences are neeeded.

Summary of the Invention

The present invention relates to methods for providing tools for functional analysis of genes defined by nucleotide sequences, for example, of cDNA and EST libraries of a genome database.

In one aspect, the invention provides a method of identifying at least one binding partner bound to a target, from a subset of binding partners directed against a plurality of discrete targets, the binding partner being a member of a larger set of binding partners. The method includes the steps of generating an array of targets, each target being affixed to a support at an addressible location; contacting the array of targets with the set of binding partners to generate a subset of binding partners bound to the targets on the

array; selecting the subset of binding partners by removing unbound binding partners; and identifying at least one of the binding partners bound to a target at an addressible location.

In preferred embodiments: the set of binding partners comprises a phage antibody library; each phage of the phage antibody library comprises an antibody selected from the group consisting of an Fv, V_H, V_L, an Fab', and an Fab; the antibody is a single chain Fv or Fab; the single chain Fv or Fab is designed from antibody framework sequences which are from a human gene or are humanized; the plurality of discrete targets is a plurality of peptides encoded by a corresponding plurality of nucleic acid sequences; the plurality of nucleic acid sequences are obtained from a genome; the plurality of encoded peptides are produced by coupled transcription and translation of the plurality of nucleic acid sequences; the plurality of encoded peptides are chemically synthesized on the support; the support is selected from the group consisting of a bead, a chip, a SELDI matrix, and a culture dish; the peptides are post-translationally modified; the peptides are glycopeptides; the step of generating the array of targets at addressible locations is monitored by SELDI.

In another aspect, the invention provides a method of determining biological characteristics of a plurality of peptides encoded by a plurality of nucleic acid sequences. The method includes the steps of: providing a plurality of distinct nucleic acid sequences; generating an array of peptide targets encoded by the distinct nucleic acid sequences, each target of the array being affixed to a support at an addressible location; contacting the array of peptide targets with a set of antibodies to generate a subset of antibodies bound to the targets; and identifying at least one antibody bound to a target in an addressible location.

In preferred embodiments: the antibodies are associated with encoding nucleic acids; the nucleic acids are encapsulated in virions; the peptide targets are chemically synthesized; the peptide targets has a hook; the peptide targets are produced *in vitro* by coupled transcription and translation; the nucleic acid sequences are from a cDNA library; the nucleic acid sequences are expressed sequence tags; the method further comprises the step of probing biological material with the selected antibodies to determine a specific site of binding.

Detailed Description of the Invention

The present invention relates to methods for producing reagents which can be used to identify the cellular location of proteins, and, in certain embodiments, relates to methods for producing reagents which bind to a protein (or proteins) encoded by a nucleic acid sequence (or sequences) for which no function is known. For example, the

method of the invention can be applied to nucleic acid sequences from an EST database, to use the information contained in an array of ESTs representing novel DNA sequence information into a corresponding array of tools that are immediately applicable to analysis of the function of the gene represented by the EST. Such tools are applicable also to protein purification and intracellular gene therapies.

Definitions

For clarity, certain terms used throughout the specification and the claims are defined.

A "target", as used herein, refers to a peptidic moiety (e.g., a peptide, a protein, a glycoprotein, a peptidomimetic, and the like) to which a binding partner can bind. A target can be isolated from a cell or organism, or, more preferably, can be synthesized by chemical or biochemical techniques, e.g., as described herein. In a prefered ambodiment, a target is a peptide which is encoded by a naturally-occurring nucleic acid of an organism. Targets suitable for use in the methods of the invention include post-translationally modified peptides. In certain embodiments, a target can include non-naturally-occurring amino acids (such as D-amino acids) or amino acid mimetics such as are known in the art. Also included as targets are peptidic moieties which include isosteres. The term "isostere" as used herein refers to a sequence of two or more residues that can be substituted for a second sequence because the steric conformation of the first sequence fits a binding site specific for the second sequence. The term specifically includes peptide back-bone modifications (i.e., amide bond mimetics) well known to those skilled in the art.

A "discrete target" refers to a spatially isolated target having a defined structure. For example, a peptide which is chemically synthesized at a location of a chip is a discrete target. Similarly, a peptide which is produced in a well of a multi-well plate, e.g., by coupled in vitro transcription/translation of a selected nucleic acid sequence, is a discrete target. In preferred embodiments, a discrete target comprises a peptidic target which is substantially free of other peptides.

An "array" of targets comprises a plurality of discrete targets, each located at a separate addressible location on a support.

The term "binding partner," as used herein, refers to a moiety which potentially can bind with an affinity for a discrete target. It will be understood that a particular binding partner may have high affinity, moderate affinity, low affinity, or essentially no affinity for a particular peptidic target or for any target in an array of targets. Similarly, a particular target in an array of targets may or may not be bound by any binding partner in a set of binding partners, although in the methods of the invention at least one (more

preferably a plurality) of targets in an array is bound by at least one binding partner. A binding partner preferably comprises a portion which binds to a target, and a portion which encodes the binding domain. In a preferred embodiment, the binding domain is an antibody and the encoding portion is a nucleic acid. In a particularly preferred embodiment, the binding partner is an antibody which is expressed on the surface of a phage, and which is encoded by a nucleic acid of the phage.

A "set" of binding partners refers to a plurality of different binding partners, e.g., a library of binding partners such as antibodies. A set of binding partners preferably includes at least about 5 different ("diverse") binding partners, more preferably at least about 10, 20, 50, 100, 1000, 10^4, 10^5, 10^6, 10^7, 10^8, 10^9 or 10^{10} binding partners or more. A"subset" refers to a portion of a set of binding partners. A subset which is "directed against a plurality of targets is a subset of binding partners which bind to that plurality of targets.

The term "plurality" as used herein, means more than one, preferably at least 2, at least 10, at least 20, at least 50, or at least 200.

Methods

In one aspect, the invention provides methods for creating reagents which are useful, e.g., for the assignment of biological characteristics to nucleic acid sequences taken from gene fragments. The invention relates to a method of identifying reagents for targets such as the protein products of ESTs from a genome, including a human genome. For example, EST sequences which can be used in the present invention can be obtained from GenBank, or other public repositories of the human genome project. The invention preferably provides methods yielding reagents to bind specifically to the peptide target within a plurality of different peptide targets, each at an addressible location in an array.

In one embodiment, the invention provides a method of identifying at least one binding partner bound to a target, from a subset of binding partners directed against a plurality of discrete targets, in which the at least one binding partner bound to a target is a member of a larger set of binding partners. The method comprising the steps of generating an array of targets, each target being affixed to a support at an addressible location; contacting the array of targets with the set of binding partners to generate a subset of binding partners bound to the targets on the array; selecting the subset of binding partners by removing unbound binding partners; and identifying at least one of the binding partners bound to a target at an addressible location.

In certain embodiments, each target of the array of discrete targets is a peptide which is encoded by a discrete nucleic acid sequence. A peptide sequence encoded by a nucleic acid can be prepared by chemical or biochemical methods. For example, the

peptides encoded by a plurality of nucleic acid sequences (illustratively from a library of DNA sequences) can be chemically synthesized on a solid support according to methods known in the art, for example, the light-directed solid-phase synthesis technology described in, e.g., U.S. Patent No. 5,143,854 to Pirrung et al. Arrays of peptides to which an encoding nucleic acid is attached have also been described (see, e.g., Needels, M.C. et al., Proc. Natl. Acad. Sci. U.S.A. (1993) 90(22), 10700-4). Alternatively, nucleic acids can be biochemically translated (or transcribed and then translated) *in vitro* according to known techniques, e.g., with the use of ribosome or wheat-germ derived kits which are commercially available for this purpose. In this embodiment, certain post-translational modifications can be provided on the targets by addition of enzymes, such as glycosidases, which cause post-translation modification of the corresponding naturally-occurring peptide in a host organism.

Databases of partial cDNA species, known as expressed sequence tags (ESTs), are growing in size (see, e.g., a description of human brain ESTs, in Adams, M. *et al.* 1993 Nature Genetics 4: 256-267). A collation of 174,472 ESTs comprising more than 52 million nucleotides of human DNA sequence has been reported (Venter, J.C. 1997 Prot. Eng. 10:79-80), generated from 3,300 cDNA libraries from 37 distinct organs and tissues. These ESTs, combined with over 300,000 ESTs listed in database dbEST, yielded 70,000 non-overlapping ESTs, about 20% of which characterized previously known genes, based on significant similarities to sequences in available databases. The remainder of the ESTs identified unknown genes.

Sequences of ESTs can be obtained from publicly available databases. For the nylon-filter arrays from Clontech known as Atlas Human cDNA Expression Array (Palo Alto, CA) arranged by function, described at http://www.clontech.com, sequences are described by GenBank Accession number. Organ-specific ESTs, for example 3,400 new sequences of human brain transcripts (Adams, M. *et al.* 1993 Nature Genet. 4:256-267) are commonly deposited in Genbank, with methodology for access both of sequences and of clones by accession numbers for each, provided in the reference.

The genome of eukaryotes comprises informational DNA, in which non-noninformational sequence material inserted ("intron") into a gene is removed from the primary transcript RNA (excised) to yield mRNA, which can be reverse transcribed or copied, e.g., by the enzyme reverse transcriptase, into cDNA. A library comprising the set of cDNA molecules of a simple eukaryotic organism such as a yeast, or a multicellular plant such as an *Arabadopsis* or animal such as *Caenorhabditis elegans* can be produced by one of ordinary skill in the art, or can be purchased commercially. Similarly, cDNA libraries of specific tissues, such as liver or brain, of, for example, a mammal such as mouse, pig, rat or human, can be obtained by well-known techniques,

or in some instances are commercially available (for example, from Stratagene, La Jolla, CA). Sets of cDNA library of human genes arrays arranged by function, for example, oncogenes and tumor suppressers; transcription factors and DNA binding proteins; receptors, cell surface proteins and cell adhesion proteins, can also be obtained commercially (Clontech, Palo Alto, CA). In a preferred embodiment, a nucleic acid sequence which is used to chemically or biochemcially produce a peptide target for use in the invention comprises cDNA or another nucleic acid from which introns have been removed.

In the array of targets, each target is affixed to the support at an addressible location, so that the target location (and the identity of the target and/or a binding partner which binds to the target) can be readily identified. For example, the array can be generated (e.g., chemically synthesized) in an pattern, or in topologically or spatially segragated locations of any kind, in which the identity of the discrete target at each location is known or can be determined, and which can be addressed (e.g., sampled, analyzed, washed, etc.) individually. In a preferred embodiment, an array is generated by automated means, such as a robot or a computer-controlled chemical synthesis apparatus. When an automated apparatus is employed to create the array of targets, the identity of each component of the array (i.e., at each spatially-separate location) can be stored in a computer memory for subsequent recall. Exemplary arrays include arrays of wells in a culture plate (e.g., a 96-well plate); arrays of polystyrene pins upon which discrete targets are immobilized; arrays on chips (e.g., as described in U.S. Patent No. 5,143,854); arrays of beads, e.g., resin beads; and the like. The support can be solid, e.g., glass, plastic, silicon, and the like; a support can also be, e.g., a resin bead, a magnetic microbead, and the like. A support can be rigid (such as a plastic culture dish); or flexible (e.g., a nylon filter). In certain preferred embodiments, an array can be generated on a surface suitable for use in analytical methods such as SELDI (Surfaces Enhanced for Laser Desorption/Ionization) (see, e.g., PCT Publication WO94/28418). In this embodiment, the targets are affixed to a SELDI surface, so that samplings and analyses of the targets (and/or binding partners bound to one or more targets) can be achieved with laser irradiation.

The targets thus arrayed can then be contacted with a diverse population of binding partners, so that binding partners bearing features (such as antibodies) that bind to the target are bound to the target, and subsequently can be separated from unbound members of the population. In a preferred embodiment, the array is contacted with a solution which contains a library of binding partners (such as antibodies) under conditions such that substantially all the locations on the array are contacted with a single aliquot of the binding partner solution. This has the advantage of conserving the

amount of binding partner soluion which must be used to bind the binding partners to targets on the array, and often simplifies the experimental procedure. For example, an array of targets affixed to a support can be washed with a solution which contains a library of binding partners (e.g., by immersing a plate on which an array of targets has been affixed in the solution), under conditions which permit at least one binding partner to bind to at least one target (and preferably more than one target). The population of diverse binding partners is preferably a library of diverse antibodies or antibody fragments, preferably with each antibody associated covalently with the nucleic acid sequence which encodes the antibody chain(s). In a preferred embodiment, the population of diverse binding partners is a library of phage, e.g., bacteriophage (bacterial viruses), in which each member of the phage library expresses a single phenotype of antibody or antibody fragment on the surface of that phage. However, in other embodiments, the population is a library of cells or cell-like biological units, for example, a population of *E. coli* cells expressing a library of diverse antibodies on a porin protein; or a population of *Bacillus subtilis* or *Streptomyces lividans* spores, which express a library of diverse antibodies on a spore coat protein, preferably with an intervening linker to permit translational and rotational access to features of the antibody. Such libraries of cells or spores can be prepared according to well-known methods. In another preferred embodiment, a library of diverse antibodies can be displayed on the surface of a eukaryotic virus, for example a retrovirus or an adenovirus, which has been engineered to eliminate pathenogenicity and infectivity, and to display an encoded peptide. A library of diverse displayed peptides (e.g., antibodies) carried by the virus, phage, cell, or cell-like structure can be engineered by mutagenesis of the encoded peptide, or by insertion of diverse nucleic acids into restriction sites, by methods known to one of ordinary skill in the art. In a preferred embodiment, an antibody which is a member of a library of antibodies is a single chain Fv, combining the binding elements of the heavy and light chains of an antibody.

The ligand moiety binding partner with specific affinity for a target is a member of a subset of binding partners, members of which have affinity for targets in an addressable array of targets. The binding partners in this subset are found in a set of binding partners, which set comprises a defined region of structural diversity in a feature of the peptide (see, for example, PCT Publications WO93/06213; WO93/11236; WO93/17105; WO93/19172; WO94/06920; WO94/13804; WO94/18221; PCT/US95/11235; EP 440,146 B1; WO95/01438; EP 436,597 B1; EP 239,400; and U.S. Patent Nos. 5,395,750; 5,403,484; 5,464,745; 5, 514,548; 5,516,637; 5,521,077; 5, 565,332 5,571,698; 5,580,717; 5,591,604; and 5,627,024).

The target-binding domain of the binding partner is generally engineered to be displayed, for example, on or close to the amino terminus of a bacteriophage filamentous protein such as M13phage gene III, or *Escherichia coli* F pilin. The set of binding partners (e.g., a library of phage antibodies) preferably can contain at least 10^3-fold diversity (i.e., at least 10^3 distinct phage expressing distinct antibodies), more preferably 10^6 to 10^9-fold diversity, and diversity of 10^{13}-10^{15} can be obtained. Preferably the set of binding partners contains at least 2, more preferably 5, 10, 100, or 1,000 copies of each diverse member, to ensure that sufficient numbers of phage are present to bind to as many targets as possible.

The term "antibody" as used herein refers to molecules such as immunoglobulins and immunologically active determinants of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen. Structurally, the simplest naturally occurring antibody (e.g., IgG) comprises four polypeptide chains, two copies of a heavy (H) chain and two of a light (L) chain, all covalently linked by disulfide bonds. Specificity of binding in the large and diverse set of antibodies is found in the variable (V) determinant of the H and L chains; regions of the molecules that are primarily structural are constant (C) in this set.

Structurally, a naturally occurring antibody (e.g., IgG, M_r 150 kDa) consists of four polypeptide chains, two copies of a heavy (H) chain and two of a light (L) chain (M_r 25 kDa), the four chains being covalently linked by disulfide bonds. Specificity of antigen binding by each molecule that comprises the large and diverse set of antibodies is found in the variable (V) determinant of the H and L chains; regions of the molecules that are primarily structural are constant (C) in this set. IgA (M_r 160 kDa) is the predominant antibody in secretions (saliva, tears, mild, nasal mucus and gastrointestinal and respiratory secretions). IgA can exist as a monomer, dimer, and higher multimeric forms (J. Kendrew, Ed., The Encyclopedia of Molecular Biology, 1994, Blackwell Science, Oxford). A further isotype is IgM can be found as a membrane-bound monomer (M_r 190 kDa) on B cells, and as a circulating secretory form as a pentamer (M_r ca. 950 kDa) that differs from the bound form at the N-terminus of the heavy chain due to alternative splicing. The heavy chain of the pentamer is attached by disulfide bridges to a J chain molecule. Minor antibody isotypes include IgD (M_r 175 kDa), expressed on cell surfaces, and IgE. IgE (total M_r 190 kDa) comprises 0.0003% of serum immunoglobulin, however it can be substantially elevated in an allergic subject such as an asthmatic, and is medically important as the major mediator of immediate type hypersensitivity.

The binding sites of the proteins that comprise an antibody, i.e., the antigen-binding functions of the antibody, are localized by analysis of fragments of a naturally-occurring antibody. Thus, antigen-binding fragments are also intended to be designated by the term "antibody." Examples of binding fragments encompassed within the term antibody include: a Fab fragment consisting of the V_L, V_H, $C_L$and C_{H1} domains; an Fd fragment consisting of the V_H and C_{H1} domains; an Fv fragment consisting of the V_L and V_H domains of a single arm of an antibody; a dAb fragment (Ward *et al.*, 1989 *Nature* 341:544-546) consisting of a V_H domain; an isolated complementarity determining region (CDR); and an $F(ab')_2$ fragment, a bivalent fragment comprising two Fab' fragments linked by a disulfide bridge at the hinge region. These antibody fragments are obtained using conventional techniques well-known to those with skill in the art, and the fragments are screened for utility in the same manner as are intact antibodies. The term "antibody" is further intended to include chimeric molecules having at least one antigen binding determinant derived from an antibody molecule. Furthermore, although the H and L chains of an Fv fragment are encoded by separate genes, a synthetic linker can be made that enables them to be made as a single protein chain (known as single chain antibody, sAb; Bird *et al.* 1988 *Science* 242:423-426; and Huston *et al.* 1988 *PNAS* 85:5879-5883) by recombinant methods. Such single chain antibodies are also encompassed within the term "antibody", and can be utilized as binding determinants in the design and engineering of a library of diverse binding vehicles. Antibody fragments are also useful for modulating the number of receptors for that antibody on the surface of cells, and for obtaining agents that mimic this activity, by screening for such agents in an assay for modulation of the receptor.

Polyclonal antibodies are produced by immunizing animals, usually a mammal, by multiple subcutaneous or intraperitoneal injections of an immunogen (antigen) and an adjuvant as appropriate. As an illustrative embodiment, animals are typically immunized against a protein, peptide or derivative by combining about 1 μg to 1 mg of protein capable of eliciting an immune response, along with an enhancing carrier preparation, such as Freund's complete adjuvant, or an aggregating agent such as alum, and injecting the composition intradermally at multiple sites. Animals are later boosted with at least one subsequent administration of a lower amount, as 1/5 to 1/10 the original amount of immunogen in Freund's complete adjuvant (or other suitable adjuvant) by subcutaneous injection at multiple sites. Animals are subsequently bled, serum assayed to determine the specific antibody titer, and the animals are again boosted and assayed until the titer of antibody no longer increases (i.e., plateaus).

Such populations of antibody molecules are referred to as "polyclonal" because the population comprises a large set of antibodies each of which is specific for one of the many differing epitopes found in the immunogen, and each of which is characterized by a specific affinity for that epitope. An epitope is the smallest determinant of antigenicity, which for a protein, comprises a peptide of six to eight residues in length (Berzofsky, J. and I. Berkower, (1993) in Paul, W., Ed., Fundamental Immunology, Raven Press, N.Y., p.246). Antibody affinities for the antigen range from low, e.g. 10^{-6} M, to high, e.g., 10^{-11} M. The polyclonal antibody fraction collected from mammalian serum is isolated by well known techniques, e.g. by chromatography with an affinity matrix that selectively binds immunoglobulin molecules such as protein A, to obtain the IgG fraction. To enhance the purity and specificity of the antibody, the specific antibodies may be further purified by immunoaffinity chromatography using solid phase-affixed immunogen. The antibody is contacted with the solid phase-affixed immunogen for a period of time sufficient for the immunogen to immunoreact with the antibody molecules to form a solid phase-affixed immunocomplex. Bound antibodies are eluted from the solid phase by standard techniques, such as by use of buffers of decreasing pH or increasing ionic strength, the eluted fractions are assayed, and those containing the specific antibodies are combined.

The term "monoclonal antibody" or "monoclonal antibody composition" as used herein refers to a preparation of antibody molecules of single molecular composition. A monoclonal antibody (mAb) composition displays a single binding specificity and affinity for a particular epitope. Monoclonal antibodies can be prepared using a technique which provides for the production of antibody molecules by continuous growth of cells in culture. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein (1975, *Nature* 256:495-497; see also Brown *et al.* 1981 *J. Immunol* 127:539-46; Brown *et al.*, 1980, *J Biol Chem* 255:4980-83; Yeh *et al.*, 1976, *PNAS* 76:2927-31; and Yeh *et al.*., 1982, *Int. J. Cancer* 29:269-75) and the more recent human B cell hybridoma technique (Kozbor *et al.*, 1983, *Immunol Today* 4:72), EBV-hybridoma technique (Cole *et al.*., 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96), and trioma techniques.

A monoclonal antibody can be produced by the following steps. In all procedures, an animal is immunized with an antigen such as a protein (or peptide thereof) as described above for preparation of a polyclonal antibody. The immunization is typically accomplished by administering the immunogen to an immunologically competent mammal in an immunologically effective amount, i.e., an amount sufficient to produce an immune response. Preferably, the mammal is a rodent such as a rabbit, rat or mouse. The mammal is then maintained on a booster schedule for a time period

sufficient for the mammal to generate high affinity antibody molecules as described. A suspension of antibody-producing cells is removed from each immunized mammal secreting the desired antibody. After a sufficient time to generate high affinity antibodies, the animal (e.g., mouse) is sacrificed and antibody-producing lymphocytes are obtained from one or more of the lymph nodes, spleens and peripheral blood. Spleen cells are preferred, and can be mechanically separated into individual cells in a physiological medium using methods well known to one of skill in the art. The antibody-producing cells are immortalized by fusion to cells of a mouse myeloma line. Mouse lymphocytes give a high percentage of stable fusions with mouse homologous myelomas, however rat, rabbit and frog somatic cells can also be used. Spleen cells of the desired antibody-producing animals are immortalized by fusing with myeloma cells, generally in the presence of a fusing agent such as polyethylene glycol. Any of a number of myeloma cell lines suitable as a fusion partner are used with to standard techniques, for example, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines, available from the American Type Culture Collection (ATCC), Rockville, Md.

The fusion-product cells, which include the desired hybridomas, are cultured in selective medium such as HAT medium, designed to eliminate unfused parental myeloma or lymphocyte or spleen cells. Hybridoma cells are selected and are grown under limiting dilution conditions to obtain isolated clones. The supernatants of each clonal hybridoma is screened for production of antibody of desired specificity and affinity, e.g., by immunoassay techniques to determine the desired antigen such as that used for immunization. Monoclonal antibody is isolated from cultures of producing cells by conventional methods, such as ammonium sulfate precipitation, ion exchange chromatography, and affinity chromatography (Zola *et al.*, Monoclonal Hybridoma Antibodies: Techniques And Applications, Hurell (ed.), pp. 51-52, CRC Press, 1982). Hybridomas produced according to these methods can be propagated in culture *in vitro* or *in vivo* (in ascites fluid) using techniques well known to those with skill in the art.

For therapeutic use of antibodies of non-human origin in humans, the non-human "foreign" epitopes elicit immune response in the patient. If sufficiently developed, a potentially lethal disease known as HAMA (human antibodies against mouse antibody) may result. To eliminate or minimize HAMA, it is desirable to engineer chimeric antibody derivatives, i.e., "humanized" antibody molecules that combine the non-human Fab variable region binding determinants with a human constant region (Fc). Such antibodies are characterized by equivalent antigen specificity and affinity of monoclonal and polyclonal antibodies described above, and are less immunogenic when administered to humans, and therefore more likely to be tolerated by the patient.

Chimeric mouse-human monoclonal antibodies (i.e., chimeric antibodies) can be produced by recombinant DNA techniques known in the art. For example, a gene encoding the Fc constant region of a murine (or other species) monoclonal antibody molecule is digested with restriction enzymes to remove the region encoding the murine Fc, and the equivalent portion of a gene encoding a human Fc constant region is substituted. (see Robinson *et al.*, International Patent Publication PCT/US86/02269; Akira, et al., European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al., European Patent Application 173,494; Neuberger et al., International Application WO 86/01533; Cabilly et al. U.S. Patent No. 4,816,567; Cabilly et al., European Patent Application 125,023; Better *et al.* 1988 *Science* 240:1041-1043); Liu *et al.* 1987 *PNAS* 84:3439-3443; Liu *et al.*, 1987, *J. Immunol.* 139:3521-3526; Sun *et al.* 1987 *PNAS* 84:214-218; Nishimura *et al.*., 1987, *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.*, 1988, *J. Natl Cancer Inst.* 80:1553-1559.)

An antibody can be humanized by any method, which is capable of replacing at least a portion of a CDR of a human antibody with a CDR derived from a non-human antibody. Winter describes a method which may be used to prepare the humanized antibodies of the present invention (UK Patent Application GB 2188638A, filed on March 26, 1987), the contents of which is expressly incorporated by reference. The human CDRs may be replaced with non-human CDRs using oligonucleotide site-directed mutagenesis as described in International Application WO 94/10332 entitled, *Humanized Antibodies to Fc Receptors for Immunoglobulin G on Human Mononuclear Phagocytes.*

The chimeric antibody can be further humanized by replacing sequences of the Fv variable region which are not directly involved in antigen binding with equivalent sequences from human Fv variable regions. General reviews of humanized chimeric antibodies are provided by Morrison, S. L., 1985, *Science* 229:1202-1207 and by Oi *et al.*, 1986, *BioTechniques* 4:214. Those methods include isolating, manipulating, and expressing the nucleic acid sequences that encode all or part of immunoglobulin Fv variable regions from at least one of a heavy or light chain. Sources of such nucleic acid are well known to those skilled in the art and, for example, may be obtained from 7E3, an anti-$GPII_bIII_a$ antibody producing hybridoma. The recombinant DNA encoding the chimeric antibody, or fragment thereof, can then be cloned into an appropriate expression vector. Suitable humanized antibodies can alternatively be produced by CDR substitution U.S. Patent 5,225,539; Jones *et al.* 1986 *Nature* 321:552-525; Verhoeyan *et al.*. 1988 *Science* 239:1534; and Beidler *et al.* 1988 *J. Immunol.* 141:4053-4060).

Human mAb antibodies directed against human proteins can be generated using transgenic mice carrying the complete human immune system rather than the mouse system. Splenocytes from these transgenic mice immunized with the antigen of interest are used to produce hybridomas that secrete human mAbs with specific affinities for epitopes from a human protein (see, e.g., Wood *et al.* International Application WO 91/00906, Kucherlapati *et al.* PCT publication WO 91/10741; Lonberg *et al.* International Application WO 92/03918; Kay *et al.* International Application 92/03917; Lonberg, N. *et al.* 1994 *Nature* 368:856-859; Green, L.L. *et al.* 1994 *Nature Genet.* 7:13-21; Morrison, S.L. *et al.* 1994 *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Bruggeman *et al.*. 1993 *Year Immunol* 7:33-40; Tuaillon *et al.* 1993 *PNAS* 90:3720-3724; Bruggeman *et al.*. 1991 *Eur J Immunol* 21:1323-1326).

Monoclonal antibodies can also be generated by other methods well known to those skilled in the art of recombinant DNA technology. An alternative method, referred to as the "combinatorial antibody display" method, has been developed to identify and isolate antibody fragments having a particular antigen specificity, and can be utilized to produce monoclonal antibodies (for descriptions of combinatorial antibody display see e.g., Sastry *et al.* 1989 *PNAS* 86:5728; Huse *et al.* 1989 *Science* 246:1275; and Orlandi *et al.* 1989 *PNAS* 86:3833). After immunizing an animal with an immunogen as described above, the antibody repertoire of the resulting B-cell pool is cloned. Methods are generally known for obtaining the DNA sequence of the variable regions of a diverse population of immunoglobulin molecules by using a mixture of oligomer primers and PCR. For instance, mixed oligonucleotide primers corresponding to the 5' leader (signal peptide) sequences and/or framework 1 (FR1) sequences, as well as primer to a conserved 3' constant region primer can be used for PCR amplification of the heavy and light chain variable regions from a number of murine antibodies (Larrick *et al.*,1991, *Biotechniques* 11:152-156). A similar strategy can also been used to amplify human heavy and light chain variable regions from human antibodies (Larrick *et al.*, 1991, *Methods: Companion to Methods in Enzymology* 2:106-110).

Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP*™ *Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* International Publication No. WO 92/18619; Dower *et al.* International Publication No. WO 91/17271; Winter *et al.* International Publication WO 92/20791; Markland *et al.* International Publication No. WO 92/15679; Breitling *et al.* International Publication WO 93/01288; McCafferty *et al.*

International Publication No. WO 92/01047; Garrard *et al.* International Publication No. WO 92/09690; Ladner *et al.* International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum Antibod Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J Mol Biol* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *PNAS* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc Acid Res* 19:4133-4137; Barbas *et al.* (1991) *PNAS* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

Specific binding proteins with high affinities for targets can be made according to methods known to those in the art. For example, proteins that bind specific DNA sequences may be engineered, and proteins that bind a variety of targets, especially protein targets (Ladner, R.C., *et al.*, U.S. Patent 5,233,409; Ladner, R.C., *et al.*, U.S. Patent 5,403,484) may be engineered and used in the present invention as the FcαR binding determinant or as the target binding determinant, as appropriate. Further, the methods of these libraries can be used in screens to obtain binding determinants that are mimetics of the structural determinants of antibodies.

In particular, the Fv binding surface of a particular antibody molecule interacts with its target ligand according to principles of protein-protein interactions, hence sequence data for V_H and V_L (the latter of which may be of the κ or λ chain type) is the basis for protein engineering techniques known to those with skill in the art. Details of the protein surface that comprises the binding determinants can be obtained from antibody sequence information, by a modeling procedure using previously determined three-dimensional structures from other antibodies obtained from NMR studies or crytallographic data. See for example Bajorath, J. and S. Sheriff, 1996, *Proteins: Struct., Funct., and Genet. 24 (2)*, 152-157; Webster, D.M. and A. R. Rees, 1995, "Molecular modeling of antibody-combining sites," in S. Paul, Ed., *Methods in Molecular Biol. 51*, Antibody Engineering Protocols, Humana Press, Totowa, NJ, pp 17-49; and Johnson, G., Wu, T.T. and E.A. Kabat, 1995, "Seqhunt: A program to screen aligned nucleotide and amino acid sequences," in *Methods in Molecular Biol.51, op. cit.*, pp 1-15.

Equivalents

The step of contacting the array of targets with the set of binding partners to generate a subset of binding partners bound to the targets on the array should be performed under conditions which permit the binding of some (but preferably not all) of the binding partners to at least one, and preferably more than one, of the immobilized

peptidic targets. However, some targets in the array of targets may not be bound by any binding partner. Conditions for promoting the binding of a binding partner to a target are well known in the art. The particular conditions employed can be varied to permit only binding partners which bind to a target with a pre-selected affinity to become bound to or associated with a target. In preferred embodiments, a binding partner is an antibody which binds to a selected target with an affinity of at least about 10^5, more preferably 10^6, more preferably 10^7, more preferably 10^8, and most preferably at least about 10^9 $mole^{-1}$.

Once a subset of binding partners bound to the targets on the array has been generated, the bound subset of partners is separated from the unbound binding partners which remain in solution. The unbound binding partners can be removed, e.g., by one or more washings with solutions under conditions which will remove unbound binding partners while permitting binding partners of a selected affinity remain bound to their respective targets. For example, the array can be repeatedly washed with buffers of increasing proton concentration (decreasing pH), increasing urea concentration, or increasing detergent concentration, or combinations thereof. Howver, in preferred embodiments, the contacting buffers are selected to avoid the inactivation of bound binding partners, including the target-binding region and the coding region of the binding partner. Washing steps can be monitored to determine the extent to which binding partners (e.g., phage or cells) are released at each washing step. For example, wash fluids can be collected and analyzed to determine the extent of release of binding partners. Alternatively, or in addition, the degree to which binding partners remain bound to the array can be monitored (e.g., by SELDI). Washing of individual locations of the array can be performed to determine whether a particular target has been bound by a binding partner; this is facilitated, e.g., when the array is an array of wells in a culture plate. Washing can be halted at any time to permit the release or retention of binding partners having a pre-selected affinity for a target in the array of targets.

Once the subset of binding partners bound to targets in the array has been formed, and the remainder of the (unbound) set of binding partners has been removed, the bound binding partners can be identified. For example, when antibodies of a phage antibody library are employed as the set of binding partners, phage which remain bound to a target can be eluted and used to infect a host (e.g., a bacterium); the phage can then be grown and sequenced to determine the identity of the antibody which bound to target. The nucleic acid sequence which encodes the antibody can be cloned to provide large quantities of such an antibody for use, e.g., in structural or functional studies as are described herein. Elution and further processing of phage can be achieved using known techniques. For example, a quantity of a sample of the specifically bound phage,

released from the array of targets, can be contacted with freshly grown host *E. coli* cells, and then contacted with soft nutrient agar. the infected bacteria are then deposited on nutrient agar in a Petri dish, to recover plaques consisting of areas of cycles of phage growth and reinfection of the proximate cells, within an area or "lawn" of uninfected cells. Individual plaques can be sampled and the antibody-encoding nucleic acid portion of the phage can be sequenced, to determine the size of the subset of bound phage. If the subset is large and diverse, a quantity of the subset that includes several different members can be amplified, for example by another cycle of contacting the phage with the target and selection of bound phage. The amplification process can be iterated, for example, two or three times, generally until a subset of phage with some (possibly complete) homogeneity of nucleotide sequence is obtained. Each selected phage will include a nucleic acid sequence which codes for the amino acid sequence of the peptide chains of the antibody displayed on that phage. In this manner, a plurality of antibodies, each specific for one of a plurality of targets is obtained. When the peptidic targets are each encoded by an EST, e.g., of the human genome, the method can provide antibodies which bind to the peptide products of the ESTs, even when the function of the EST products is unknown.

In certain embodiments, the targets are not immobilized prior to contacting with the binding partner (e.g., antibody) library, but are capable of being immobilized to a support after contact with the binding partner library (and binding of binding partners to at least one target). For example, a plurality of different targets can be contacted in solution by a library of diverse binding partners; following binding of the binding partners to the targets, the complexes of targets with their respective binding partners can be immobilized in an addressible array. Following immobilization, the immobilized array can be washed to remove unbound binding partners (e.g., as described elsewhere herein) to select specifically bound partners. A peptidic target (and in certain cases an associated binding partner) can be immobilized on a support by including a"hook" in the peptidic target (or in the binding partner) which permits selective binding of the target to the support. For example, a peptide (or a plurality of peptides) can be chemically synthesized to include an amino acid which has been modified to include a member of a specific binding pair, e.g., biotin (the "hook"). A solid support which is provided with a bound complement to the specific binding pair (e.g., avidin or streptavidin to bind a biotinylated target) can then bind the target. It will be appreciated that the peptides must be spatially segaegated from each other prior to immobilization on the support, e.g., the peptides can be synthesized, and then immobilized, in individual wells of a multi-well plate, to form (e.g., after contacting with a library of binding partners) an array of immobilized peptidic targets. In another embodiment, the peptide target includes a

sequence of six histidine residues at a terminus of the peptide, and the solid support contains nickel ions, so that the peptide is coupled by affinity to the solid support. The hexahistidine addition is a "hook" by which the plurality of targets can be affixed to solid support. Hooks for the peptides can be based on non-covalent affinity for the surface (for example, a hexahistidine portion of a peptide binds to a nickel-containing solid support), or the binding of a peptidic target to a solid support can be covalent. Another example of a hook is an antigenic sequence of a target which can be recognized by an antibody bound to a support. For example, a plurality of peptide targets can be chemically synthesized such that each peptide includes an identical antigenic region to which an antibody can bind. the targets can then be bound to a solid surface coated with the corresponding antibody.

In certain embodiments, the antibody displayed by the members of the diverse phage library is a variable portion of either a heavy chain or a light chain, for example, a V_H or a V_L. The displayed domain can be an Fab' or an Fab. In a preferred embodiment, the single chain Fv or Fab displayed on the phage or cell is designed from antibody framework sequences which are from a human gene or are humanized.

In another aspect, the invention provides a method of determining biological characteristics of a plurality of peptides encoded by a plurality of nucleic acid sequences. The method includes the steps of providing a plurality of distinct nucleic acid sequences; generating an array of peptide targets encoded by the distinct nucleic acid sequences, each target of the array being affixed to a support at an addressible location; contacting the array of peptide targets with a set of antibodies to generate a subset of antibodies bound to the targets; and identifying at least one antibody bound to a target in an addressible location. In a preferred embodiment, the antibodies are associated with a phage, a virus, a cell or a cell-like structure which contains the nucleic acid sequence encoding a variable feature of the antibody, one example of which is found displayed on the phage or cell or cell-like structure.

In preferred embodiments, the step of providing a plurality of distinct nucleic acid sequences comprises providing a plurality of distinct nucleic acid sequences from a database of nucleic acid sequences (e.g., a plurality of ESTs). The plurality of peptide targets can be chemically synthesized, e.g., by using genetic code to determine the identity of the peptides encoded by the plurality of EST sequences, and chemically synthesizing the plurality of peptides. In other embodiments, the targets are produced *in vitro* by translation, or in yet another embodiment, by coupled transcription and translation. In this embodiment, the peptide targets preferably contain a hook, by which the target molecules can be immobilized to a solid support.

In any of the embodiments described hereinabove, once a binding partner (e.g., an antibody) which binds to a particular target in the array has been identified, the binding partner can be used to probe biological sample (such as a cell, a tissue, an organ, a fluid sample, and like) to determine the presence or absence of the target in the biological sample. In certain embodiments, the method further comprises the step of probing biological material with the selected antibodies to determine a specific site of binding. The biological material can be a cell or set of cells from an organ, from a tissue, or it can be extracellular lamella material. The biological material can be an entire organism. The antibodies identified or made by the method of the invention are preferably labeled, e.g., with a radioisotope label, a luminescent or fluorescent label, as is conventional in the art. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. A labeled antibody can then be employed to detect, quantitate, or localize a peptidic target to which the alabeled antibody binds. For example, the organism, organ, tissue, or cells can be probed during the course of development of the organism to determine the time during development at which a target peptide is expressed. A sample can be probed during the course of infection with a pathogenic agent to determine the distribution of peptidic targets from the pathogen in the sample, and the like. Antibodies can also be used to inhibit the activity of a target peptide, ot to inhibit the binding of other cellular comonents, such as binding proteins, to the target peptide of interest.

Preferred methods of observation and analysis include direct visualization with a microscope fitted with a light source and filters appropriate to the excitation and emission wavelengths, and use of a camera attached to the microscope. Another preferred method of analysis is cell separation and enumeration of live cells appropriately stained with this class of dye reagents, is isolation by use of a flow cytometry apparatus such as a FACScan or a FACStar (Becton Dickinson, San Jose, CA). This instrument illuminates a mixed cell population, for example at 488 nm with an argon laser source of light, and uses an emission spectrum signal from each cell detected in a moving fluid such as a buffer, to sort each cell as it is flowing past the detector using a variety of bandpass filters for collection of emitted light (see, for example, Lohmeyer, J. *et al.*, 1994, *J. Immunol. Methods 172*:59). The apparatus can

count and collect cell populations yielding both data and cell fractions for further analysis and use.

Another fluorescent dye component of numerous antibody conjugate reagents is phycoerythrin, an accessory photosynthetic pigment found in red algae, which functions *in vivo* to transfer light energy to chlorophyll during photosynthesis. The 240 kDA protein has 34 phycoerythrobilin fluorochromes, each a linear tetrapyrrole, per molecule, which when excited by 488-nm light, emit light at a 576 nm peak. For single-laser flow cytometer use, a 585 +/-21 nm BP filter is used for optimal detection. When performing multi-color analysis with a dual-laser system, a tighter window of detection is required to compensate for the other conjugates being used, for example a 575 +/-13 nm BP filter. Conjugation chemistry for PharMingen (San Diego, CA) PE products yields an average of one PE molecule per antibody or other protein, for example a binding partner or an antibody for a target produced by the methods of the present invention.

A binding partner produced by the methods of the invention, bound to a target in biological material, can be visualized by use of covalently-linked bioluminescent protein, e.g., a luciferase, for example a firefly luciferase (*Photinus*) or a coelenterate luciferase (*Renilla*). Methods and reagents are described, for example, in U.S. Patent Nos. 5,292,658, 5,035,999, and 5,491,084.

Molecules produced by the methods of the invention can be labeled also by genetic fusion to an enzyme active site, for example, to alkaline phosphatase or β-galactosidase. Extend of binding to a target in a biological material can be determined by colorimetric means, by provision of a chromogenic substrate (X-P or X-gal, wherein X indicates (5-bromo- 4-chloro- 3-indolyl-) for -phosphate, or -β-D- galactoside, respectively). Enzymatically labeled reagents can be used to determine target location and developmental stage location within the gross anatomy, for example, organ-specific expression and embryonic, neonatal, juvenile or adult onset of expression.

Equivalents

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, that the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications within the scope of the invention will be apparent to those skilled in the art to which the invention pertains.

All contents of all patents and publications referred to herein are hereby incorporated by reference.

What is claimed is:

1. A method of identifying at least one binding partner bound to a target, from a subset of binding partners directed against a plurality of discrete targets, the binding partner being a member of a larger set of binding partners, the method comprising the steps of:

generating an array of targets, each target being affixed to a support at an addressible location;

contacting the array of targets with the set of binding partners to generate a subset
of binding partners bound to the targets on the array;

selecting the subset of binding partners by removing unbound binding partners; and

identifying at least one of the binding partners bound to a target at an addressible location.

2. The method of claim 1, wherein the set of binding partners comprises a phage antibody library.

3. The method of claim 2, wherein each phage of the phage antibody library comprises an antibody selected from the group consisting of an Fv, V_H, V_L, an Fab', and an Fab.

4. The method of claim 3, wherein the antibody is an Fv or Fab.

5. The method of claim 4, wherein the antibody is a single chain Fv or Fab.

6. The method of claim 5, wherein the single chain Fv or Fab is designed from antibody framework sequences which are from a human gene or are humanized.

7. The method of claim 1, wherein the plurality of discrete targets is a plurality of peptides encoded by a corresponding plurality of nucleic acid sequences.

8. The method of claim 7, wherein the plurality of nucleic acid sequences are obtained from a genome.

9. The method of claim 8, wherein the plurality of encoded peptides are produced by coupled transcription and translation of the plurality of nucleic acid sequences.

10. The method of claim 8, wherein the plurality of encoded peptides are chemically synthesized on the support.

11. The method of claim 1, wherein the support selected from the group consisting of a bead, a chip, a SELDI matrix, and a culture dish.

12. The method of claim 9, wherein the peptides are post-translationally modified.

13. The method of claim 12, wherein the peptides are glycopeptides.

14. The method of claim 1, wherein the step of generating the array of targets at addressible locations is monitored by SELDI.

15. A method of determining biological characteristics of a plurality of peptides encoded by a plurality of nucleic acid sequences, the method comprising the steps of:

providing a plurality of distinct nucleic acid sequences;

generating an array of peptide targets encoded by the distinct nucleic acid sequences, each target of the array being affixed to a support at an addressible location;

contacting the array of peptide targets with a set of antibodies to generate a subset of antibodies bound to the targets; and

identifying at least one antibody bound to a target in an addressible location.

16. The method of claim 15, wherein the antibodies are associated with encoding nucleic acids.

17. The method of claim 16, wherein the nucleic acids are encapsulated in virions.

18. The method of claim 15, wherein the peptide targets being chemically synthesized.

19. The method of claim 15, wherein at least one of the peptide targets has a hook.

20. The method of claim 15, wherein the peptide targets are produced *in vitro* by coupled transcription and translation.

21. The method of claim 15, wherein the nucleic acid sequences are from a cDNA library.

22. The method of claim 15, wherein the nucleic acid sequences are expressed sequence tags.

23. The method of claim 15, wherein the method further comprises the step of probing biological material with the selected antibodies to determine a specific site of binding.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/15895

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :G01N 33/53; C12Q 1/68
US CL : 435/7.1, 6, 5, 7.92; 436/518, 527, 524
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.1, 6, 5, 7.92; 436/518, 527, 524

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAS, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,527,681 A (HOLMES) 18 June 1996, col. 2, lines 1-41; col. 7, lines 44-60; col. 9, lines 21-32; col. 10, lines 4-10; col. 16, last paragraph.	1, 8, 9, 11, 14
Y	US 5,464,745 A (MIERENDORF et al) 07 November 1995, col. 4, lines 3-24; col. 6, lines 50-60.	1-23
A	JANDA, K. D. Tagged versus untagged libraries: Methods for the generation and screening of combinatorial chemical libraries. Proc. Natl. Acad. Sci. (USA). November 1994, Vol. 91, pages 10799-10785, especially page 10782.	1-23
Y	US 5,514,548 A (KREBBER et al) 07 May 1996, col. 1,line 40 to col. 4, line 41.	1-12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 30 SEPTEMBER 1998	Date of mailing of the international search report 30 OCT 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer T. WESSENDORF Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/15895

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,521,077 A (KHOSLA et al) 28 May 1996, entire document.	1-23
A	US 5,143,854 A (PIRRUNG et al) 01 September 1992, entire document.	1-23